

Description**Method and arrangement for transforming a picture area**

- 5 The invention relates to a method and an arrangement for transforming a picture area.

Such a method with an associated arrangement is disclosed in [1]. The known method serves in the MPEG  
10 standard as a coding method and is essentially based on the hybrid DCT (Discrete Cosine Transform) with motion compensation. A similar method is used for videophony at  $n \times 64$  kbit/s (CCITT Recommendation H.261), for TV contribution (CCR Recommendation 723) at 34 or  
15 45 Mbit/s, and for multimedia applications at 1.2 Mbit/s (ISO-MPEG-1). Hybrid DCT comprises a temporal processing stage, which uses the relationships between successive pictures, and a spatial processing stage, which utilizes the correlation within a picture.

20 The spatial processing (intraframe coding) essentially corresponds to traditional DCT coding. The picture is broken down into blocks of  $8 \times 8$  pixels which are each transformed into the frequency domain by means of DCT.  
25 The result is a matrix of  $8 \times 8$  coefficients which approximately reflect the two-dimensional spatial frequencies in the transformed picture block. A coefficient with frequency 0 (DC component) represents and average gray-scale value of the picture block.

30 The transformation is followed by data expansion. However, in natural picture originals, a concentration of the energy around the DC component (DC value) will take place, while the very high-frequency coefficients  
35 are usually zero.

09889934-072501

In a next step, spectral weighting of the coefficients is effected, with the result that the amplitude accuracy of the high-frequency coefficients is reduced. The properties of the human eye, whereby high spatial  
5 frequencies are resolved less accurately than low spatial frequencies, are exploited in this case.

A second step of data reduction takes place in the form of an adaptive quantization through which the amplitude  
10 accuracy of the coefficients is reduced further or through which the small amplitudes are set to zero. In this case, the measure of the quantization depends on the occupancy of the output buffer: with the buffer empty, fine quantization is effected, with the result  
15 that more data are generated, while with the buffer full, coarser quantization is effected, as a result of which the volume of data is reduced.

After the quantization, the block is scanned diagonally  
20 ("zigzag" scanning), followed by entropy coding, which brings about the actual data reduction. Two effects are exploited for this purpose:

- 1.) The statistics of the amplitude values (high  
25 amplitude values occur more rarely than low ones, so that the rare events are assigned long code words and the frequent events are assigned short code words (Variable Length Coding, VLC). This results, on average, in a lower data rate than in  
30 the case of coding with a fixed word length. The variable rate of the VLC is subsequently smoothed in the buffer memory.
- 2.) Use is made of the fact that, starting from a  
35 specific value, in most cases only zeros will follow. Instead of all these zeros, only an EOB code (End Of Block) is transmitted, which leads to

105210-4E68860

a significant coding gain in the compression of the picture data. Instead of the initial rate of 512 bits, in the example specified only 46 bits need be transmitted for this

5

09889934.072501

block, which corresponds to a compression factor of more than 11.

A further compression gain is obtained through the  
5 temporal processing (interframe coding). A lower data rate is required for coding differential pictures than for the original pictures, because the amplitude values are much lower.

10 However, the temporal differences are only small if the movements in the picture are also small. By contrast, if the movements in the picture are large, then large differences are produced, which are in turn difficult to code. For this reason, the picture-to-picture motion  
15 is measured (motion estimation) and compensated (motion compensation) before the difference formation. In this case, the motion information is transmitted with the picture information, usually only one motion vector being used per macroblock (e.g. four 8x8 picture  
20 blocks).

Even smaller amplitude values of the differential pictures are obtained if motion-compensated bidirectional prediction is used instead of the  
25 prediction that is used.

In a motion-compensated hybrid coder, the picture signal itself is not transformed, but rather the temporal differential signal. For this reason, the  
30 coder is also provided with a temporal recursion loop, because the predictor must calculate the predicted value from the values of the already transmitted (coded) pictures. An identical temporal recursion loop is situated in the decoder, so that coder and decoder  
35 are fully synchronized.

0988994-07501  
T05220-4E68860

In the MPEG-2 coding method, there are principally three different methods which can be used to process pictures:

5

09080934.072501

- 5 I pictures: In the case of the I pictures, temporal prediction is not used, i.e. the picture values are directly transformed and coded, as illustrated in Figure 1. I pictures are used in order to be able to begin the decoding operation anew without knowledge of the temporal past, or in order to achieve resynchronization in the event of transmission errors.
- 10 P pictures: The P pictures are used to perform a temporal prediction; the DCT is applied to the temporal prediction error.
- 15 B pictures: In the case of the B pictures, the temporal bidirectional prediction error is calculated and then transformed. In principle, the bidirectional prediction works adaptively, i.e. forward prediction, backward prediction or
- 20 interpolation are permitted.

In MPEG-2 coding, a picture sequence is divided into so-called GOPs (Group Of Pictures). n pictures between two I pictures form a GOP. The distance between the P pictures is designated by m, in each case m-1 B pictures being situated between the P pictures. However, the MPEG syntax leaves it to the user to choose m and n. m=1 means that no B pictures are used, and n=1 means that only I pictures are coded.

A column-by-column or row-by-row transformation is preferably effected in the context of the DCT transformation on the part of the encoder. In this case, the type of transformation is effected identically for all the picture data, which is disadvantageous for specific picture data.

09889934.072501

The **object** of the invention consists in transforming a picture area, the order of vertical and

09889934-072501  
T05240-4E6B8B60

horizontal transformation depending on predetermined conditions which are taken into account in a targeted manner.

- 5 In this case, it is possible to achieve a significant improvement in the picture quality.

This object is achieved in accordance with the features of the independent patent claims. Developments of the  
10 invention also emerge from the dependent claims.

In order to achieve the object, a method for transforming a picture area is specified, in which firstly a vertical transformation of the picture area  
15 and then a horizontal transformation of the picture area or, conversely, firstly the horizontal transformation and then the vertical transformation are carried out by a decision unit.

- 20 A development consists in the picture area having an irregular structure.

In this case, it is particularly advantageous that the order of the transformations can be determined  
25 depending on a prescribed or a determined value in the decision unit or by the decision unit. Thus, depending on the picture area to be transformed and special features that are characteristic of said picture area, the order of horizontal and vertical transformation can  
30 be prescribed by the decision unit in such a way that the best possible result is obtained with regard to the compression of the picture area.

The order of the transformations is crucial in  
35 particular in the case of an irregular structure of the picture area, since, after each vertical or horizontal transformation, pixels of the irregular picture area are resorted and, as a result, a

09889934.072501



correlation of the pixels in the space domain can be lost. Such resorting may, in particular, be orientation along a horizontal or a vertical axis (line).

- 5 The decision unit determines the order of the transformations preferably using special features or a special feature of the picture area, its transmission type or a feature that is characteristic of it.
- 10 A refinement consists in the orientation of the picture area being effected along a horizontal line, or in the orientation being effected along a vertical line. In this case, pixels of the lines of the picture area are oriented on the vertical line, or pixels of the columns
- 15 of the picture area are oriented on the horizontal line. In particular, each transformation (vertical or horizontal) is followed by a corresponding orientation. As a result of the orientation, i.e. the displacement of lines and/or columns of the picture area, a
- 20 correlation in the space domain is lost under certain circumstances (in the case of an irregular structure for the picture area), since pixels originally lying next to one another will no longer necessarily lie next to one another after the orientation (e.g. correlation
- 25 in the space domain). This information is used, in particular, to take the decision about the order of the transformations within the decision unit to the effect that the correlation of pixels lying next to one another in the space or time domain is optimally
- 30 utilized.

A refinement furthermore consists in at least one of the following mechanisms being taken into account by the decision unit for determining the order of vertical

35 and horizontal transformation:

09889934 "072501

- 09889934.072504
- 5 a) In the event of transmission in the line interlacing method (interlaced) only every second line of a picture is represented (and transmitted). Alternation of the respective other second lines results, in a manner staggered over time, in pictures which represent moving pictures, the lines of in each case two temporally successive pictures complementing one another to form a frame. In the decision unit, e.g. the picture header is used to determine whether such transmission in the line interlacing method is present. If a line interlacing method is present, then the horizontal transformation is carried out first and then the vertical transformation. This exploits the fact that, in the line interlacing method, only every second line is transmitted and, consequently, the correlation of pixels is higher within a line than along a column.
- 10
- 15
- 20 b) Another mechanism consists, as described above, in that transformation being carried out first along whose direction the correlation of the picture area pixels to be transformed is greater.
- 25 Another development consists in an additional dimension being taken into account in the transformation, this additional dimension being examined with regard to the correlation of the pixels in the additional dimension. One example is that the additional dimension is a time axis (3D transformation).
- 30

A further refinement consists in a side information item containing the order of the transformations being generated by the decision unit. In this case, the side information item corresponds to a signal which is preferably transmitted to a receiver (decoder) and using which said receiver is able to infer the

35

information about the order of the transformations.  
This order is to be taken into account correspondingly  
during the inverse operation of decoding.

09889934.02501  
T052/0"4E68860

In the context of another development, the vertical transformation follows from the horizontal transformation in that mirroring is carried out on a 45° axis before the transformation. A horizontal  
5 transformation follows from the vertical transformation in a corresponding manner. The mirroring (virtually) interchanges the transformation order.

The method is suitable for use in a coder for  
10 compression of picture data, e.g. an MPEG picture coder. A corresponding decoder is preferably augmented by a possibility of evaluating the side information signal in order to be able to carry out the correct order of vertical and horizontal transformation (or the  
15 operation that is respectively the inverse thereof) during the decoding of the picture area.

Coder and decoder preferably operate according to an MPEG standard or according to an H.26x standard.

20 A development consists in the transformation being a DCT transformation or an IDCT transformation that is the inverse thereof.

25 Furthermore, in order to achieve the object, an arrangement for transforming a picture area is specified, having a decision unit using which a vertical transformation of the picture area and then a horizontal transformation of the picture area or,  
30 conversely, firstly the horizontal transformation and then the vertical transformation of the picture area can be carried out.

This arrangement is particularly suitable for carrying  
35 out the method according to the invention or one of its developments explained above.

09889934.072501

Exemplary embodiments of the invention are illustrated and explained below with reference to the drawings.

In the figures:

5

Figure 1 shows a sketch illustrating steps of a transformation of a picture area;

10

Figure 2 shows a sketch illustrating a decision unit and the signals/values generated therefrom;

15

Figure 3 shows a sketch illustrating a transmitter and receiver for picture compression;

Figure 4 shows a sketch illustrating a picture coder and a picture decoder in greater detail;

20

Figure 5 shows a possible instance of the decision unit in the form of a processor unit.

25

**Figure 1** illustrates steps of a transformation, in particular a DCT transformation for a predetermined picture area, which picture area has an irregular structure. A step 101 shows the irregular structure of the picture area in a line interlacing method, indicated by every second occupied line. In this case, the picture area is composed of the lines 105, 106, 107 and 108. In a step 102, the picture which is actually represented in the line interlacing method is shown, which again has the lines 105 to 108. The correlation of this picture area having an irregular structure is particularly high along the lines. Correspondingly, in the line interlacing method, firstly the lines are transformed after they have previously been oriented along a vertical line 109. The orientation results in a column-related displacement of adjacent pixels. The vertical transformation takes place in

30

35

09889934.072501

step 103. A horizontal orientation along a horizontal line 110 is carried out beforehand.

It would also be possible (additionally) to take  
5 account of a transformation along a time axis. Thus,  
step 101 can also be interpreted as a representation of  
a plurality of lines 105 to 108 or a plurality of  
picture areas 105 to 108 which are scanned along a time  
axis 111 at different instants in each case. The  
10 spatial information in the respective lines 105 to 108  
or the respective picture areas 105 to 108 is high,  
whereas lower correlations between the individual lines  
105 to 108 or picture areas 105 to 108 are given as a  
result of the scanning along the time axis 111 in the  
15 direction of the time dimension.

**Figure 2** illustrates a sketch illustrating a decision  
unit and the signals/values generated therefrom. An  
input signal or a plurality of input signals 200 are  
20 used by the decision unit 201 for determining which of  
a plurality of transformations (horizontal, vertical,  
temporal) are to be carried out in what order in order  
in each case to utilize the correlations in the space  
or time domain as well as possible, i.e. to take  
25 account of high correlations in such a way that an  
associated transformation is carried out first. The  
line interlacing method discussed in figure 1 serves as  
an example, which method is used by the decision unit  
201 to carry out the horizontal transformation before  
30 the vertical transformation. The actual transformations  
are carried out in a unit 202, in which the picture  
areas are likewise oriented. The resulting coefficients  
203 are the result of the transformation unit 202 (also  
cf. illustration in step 104). Furthermore, the  
35 decision unit 201 generates a side information item 203  
comprising the order of the transformations to be  
carried out.

09889934.072501

The arrangement illustrated in figure 2 is, in particular, part of a transmitter (coder) 301 as is shown in **figure 3**. Picture data 303, preferably in compressed form, are transmitted from the transmitter  
5 301 to a receiver (decoder) 302. The side information item 203 described in figure 2 is likewise transmitted (identified here by a connection 304) from the transmitter 301 to the receiver 302, where the side information item 304 is decoded to yield the  
10 information about the order of the transformations.

Moreover, it shall be pointed out that, in principle, there are two possibilities for carrying out the transformations: either both transformations  
15 (horizontal and vertical) are actually interchanged. This leads to a not inconsiderable complexity in programming terms. As an alternative to this, it is possible to define the order of the transformations (using the decision unit 201), the vertical  
20 transformation following from the horizontal transformation in that the picture area is mirrored at a 45° axis (top left to bottom right). The mirroring (virtually) interchanges the transformation order. The mirroring operation on the part of the receiver 302 is  
25 to be taken into account in a corresponding manner.

**Figure 4** shows a picture coder with an associated picture decoder in greater detail (block-based picture coding method in accordance with H.263 standard).

30

A video data stream to be coded, with temporally successive digitized pictures, is fed to a picture coding unit 201. The digitized pictures are subdivided into macroblocks 202, each macroblock having 16x16  
35 pixels. The macroblock 202 comprises 4 picture blocks 203, 204, 205 and 206, each picture block containing

09889934-072501

8x8 pixels which are assigned luminance values (brightness values). Furthermore, each macroblock 202 comprises two chrominance blocks 207 and 208 with chrominance values (color information, color saturation) assigned to the pixels.

The block of a picture contains a luminance value (= brightness), a first chrominance value (= hue) and a second chrominance value (= color saturation). In this case, luminance value, first chrominance value and second chrominance value are designated as color values.

The picture blocks are fed to a transform coding unit 209. In the case of differential picture coding, values to be coded of picture blocks of temporally preceding pictures are subtracted from the picture blocks that are currently to be coded; only the difference-information information 210 is fed to the transform coding unit (Discrete Cosine Transform, DCT) 209. To that end, the current macroblock 202 is communicated to a motion estimation unit 229 via a connection 234. In the transform coding unit 209, spectral coefficients 211 are formed for the picture blocks or differential picture blocks to be coded and are fed to a quantization unit 212. This quantization unit 212 corresponds to the quantization apparatus according to the invention.

Quantized spectral coefficients 213 are fed both to a scan unit 214 and to an inverse quantization unit 215 in a backward path. After a scan method, e.g. a "zigzag" scan method, entropy coding is carried out on the scanned spectral coefficients 232 in an entropy coding unit 216 provided for this purpose. The entropy-coded spectral coefficients are transmitted as coded picture data 217 via a channel, preferably a line or a radio link, to a decoder.

09889934-072501



09889934-072501

In the inverse quantization unit 215, inverse quantization of the quantized spectral coefficients 213 takes place. Spectral coefficients 218 obtained in this way are fed to an inverse transform coding unit 219

5 (Inverse Discrete Cosine Transform, IDCT). Reconstructed coding values (also differential coding values) 220 are fed to an adder 221 in the differential picture mode. The adder 221 furthermore receives coding values of a picture block which are produced from a

10 temporally preceding picture after a motion compensation that has already been carried out. Using the adder 221, reconstructed picture blocks 222 are formed and stored in a picture memory 223.

15 Chrominance values 224 of the reconstructed picture blocks 222 are fed from the picture memory 223 to a motion compensation unit 225. For brightness values 226, interpolation is effected in an interpolation unit 227 provided for this purpose. Using the interpolation,

20 the number of brightness values contained in the respective picture block is preferably doubled. All the brightness values 228 are fed both to the motion compensation unit 225 and to the motion estimation unit 229. The motion estimation unit 229 additionally

25 receives via the connection 234 the picture blocks of the macroblock (16x16 pixels) to be coded in each case. In the motion estimation unit 229, the motion estimation is effected taking account of the interpolated brightness values ("motion estimation on a

30 half-pixel basis"). Preferably, the motion estimation comprises the determination of absolute differences of the individual brightness values in the macroblock 222 that is currently to be coded and the reconstructed macroblock from the temporally preceding picture.

35 The result of the motion estimation is a motion vector 230, which expresses a spatial displacement of the

selected macroblock from the temporally preceding picture to the macroblock 202 to be coded.

Both brightness information and chrominance information  
5 related to the macroblock determined by the motion estimation unit 229 are displaced by the motion vector 230 and subtracted from the coding values of the macroblock 202 (see data path 231).

10 **Figure 5** shows a processor unit PRZE suitable for carrying out transformation and/or compression/decompression. The processor unit PRZE comprises a processor CPU, a memory SPE and an input/output interface IOS, which is utilized in various ways via an  
15 interface IFC: via a graphics interface, an output becomes visible on a monitor MON and/or is output on a printer PRT. An input is effected via a mouse MAS or a keyboard TAST. The processor unit PRZE also has a data bus BUS, which ensures the connection of a memory MEM,  
20 the processor CPU and the input/output interface IOS. Furthermore, additional components, e.g. additional memory, data storage device (hard disk) or scanner can be connected to the data bus BUS.

09889934"072501

Bibliography:

- [1] J. De Lameillieure, R. Schäfer: "MPEG-2-  
Bildcodierung für das digitale Fernsehen" [MPEG-2  
5 picture coding for digital television], Fernseh-  
und Kino-Technik, volume 48, No. 3/1994, pages 99-  
107.

0989934-072501